



SEQUENCE LISTING

<110> Punnonen, Juha
Apt, Doris
Neighbors, Margaret
Leong, Steven R.

<120> NOVEL TUMOR-ASSOCIATED ANTIGENS

<130> 0334.210US

<140> US 10/828,559

<141> 2004-04-19

<150> US 60/464,780

<151> 2003-04-22

<160> 95

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 185

<212> PRT

<213> Artificial Sequence

<220>

<223> TAg-25 fragment comprising extra-cellular domain
(ECD)

<400> 1

Arg	Ile	Lys	Pro	Glu	Gly	Ala	Leu	Gln	Asn	Asn	Asp	Gly	Leu	Tyr	Asp
1				5				10						15	
Pro	Asp	Cys	Asp	Glu	Ser	Gly	Leu	Phe	Lys	Ala	Lys	Gln	Cys	Asn	Gly
			20					25					30		
Thr	Ala	Thr	Cys	Trp	Cys	Val	Asn	Thr	Ala	Gly	Val	Arg	Arg	Thr	Asp
			35				40					45			
Lys	Asp	Thr	Glu	Ile	Thr	Cys	Ser	Glu	Arg	Val	Arg	Thr	Tyr	Trp	Ile
			50			55					60				
Ile	Ile	Glu	Leu	Lys	His	Lys	Glu	Arg	Glu	Ser	Pro	Tyr	Asp	Ser	Lys
65					70				75					80	
Ser	Leu	His	Thr	Ala	Leu	Gln	Lys	Glu	Ile	Thr	Thr	Arg	Tyr	Gln	Leu
				85				90						95	
Asp	Pro	Lys	Phe	Ile	Thr	Ser	Ile	Leu	Tyr	Glu	Asn	Asn	Val	Ile	Thr
			100					105					110		
Ile	Asp	Leu	Met	Gln	Asn	Ser	Ser	Gln	Lys	Thr	Gln	Asp	Asp	Val	Asp
			115				120					125			
Ile	Ala	Asp	Val	Ala	Tyr	Tyr	Phe	Glu	Lys	Asp	Val	Lys	Gly	Glu	Ser
			130				135				140				
Leu	Phe	His	Ser	Lys	Lys	Met	Asp	Leu	Arg	Val	Asn	Gly	Glu	Leu	Leu
145				150					155					160	
Asp	Leu	Asp	Pro	Gly	Gln	Thr	Leu	Ile	Tyr	Tyr	Val	Asp	Glu	Lys	Ala
				165				170						175	
Pro	Glu	Phe	Ser	Met	Gln	Gly	Leu	Lys							
			180					185							

<210> 2
 <211> 57
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> TAG-25 fragment comprising propeptide (PP)

<400> 2
 Gln Glu Glu Cys Val Cys Glu Asn Tyr Lys Leu Ala Val Asn Cys Phe
 1 5 10 15
 Val Asn Asn Asn Arg Glu Cys Gln Cys Thr Ser Val Gly Ala Gln Asn
 20 25 30
 Thr Val Ile Cys Ser Lys Leu Ala Ala Lys Cys Leu Val Met Lys Ala
 35 40 45
 Glu Met Asn Gly Ser Lys Leu Gly Arg
 50 55

<210> 3
 <211> 23
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> TAG-25 fragment comprising signal peptide (SP)

<400> 3
 Met Ala Pro Pro Gln Ala Leu Ala Leu Gly Leu Leu Leu Ala Ala Ala
 1 5 10 15
 Thr Ala Thr Phe Ala Ala Ala
 20

<210> 4
 <211> 265
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> TAG-25 polypeptide (which comprises signal
 peptide+propeptide+ ECD)

<400> 4
 Met Ala Pro Pro Gln Ala Leu Ala Leu Gly Leu Leu Leu Ala Ala Ala
 1 5 10 15
 Thr Ala Thr Phe Ala Ala Ala Gln Glu Cys Val Cys Glu Asn Tyr
 20 25 30
 Lys Leu Ala Val Asn Cys Phe Val Asn Asn Asn Arg Glu Cys Gln Cys
 35 40 45
 Thr Ser Val Gly Ala Gln Asn Thr Val Ile Cys Ser Lys Leu Ala Ala
 50 55 60
 Lys Cys Leu Val Met Lys Ala Glu Met Asn Gly Ser Lys Leu Gly Arg
 65 70 75 80
 Arg Ile Lys Pro Glu Gly Ala Leu Gln Asn Asn Asp Gly Leu Tyr Asp

				85					90					95			
Pro	Asp	Cys	Asp	Glu	Ser	Gly	Leu	Phe	Lys	Ala	Lys	Gln	Cys	Asn	Gly		
			100					105					110				
Thr	Ala	Thr	Cys	Trp	Cys	Val	Asn	Thr	Ala	Gly	Val	Arg	Arg	Thr	Asp		
		115					120					125					
Lys	Asp	Thr	Glu	Ile	Thr	Cys	Ser	Glu	Arg	Val	Arg	Thr	Tyr	Trp	Ile		
	130					135					140						
Ile	Ile	Glu	Leu	Lys	His	Lys	Glu	Arg	Glu	Ser	Pro	Tyr	Asp	Ser	Lys		
145					150				155						160		
Ser	Leu	His	Thr	Ala	Leu	Gln	Lys	Glu	Ile	Thr	Thr	Arg	Tyr	Gln	Leu		
			165					170						175			
Asp	Pro	Lys	Phe	Ile	Thr	Ser	Ile	Leu	Tyr	Glu	Asn	Asn	Val	Ile	Thr		
		180						185					190				
Ile	Asp	Leu	Met	Gln	Asn	Ser	Ser	Gln	Lys	Thr	Gln	Asp	Asp	Val	Asp		
	195						200					205					
Ile	Ala	Asp	Val	Ala	Tyr	Tyr	Phe	Glu	Lys	Asp	Val	Lys	Gly	Glu	Ser		
	210					215				220							
Leu	Phe	His	Ser	Lys	Lys	Met	Asp	Leu	Arg	Val	Asn	Gly	Glu	Leu	Leu		
225					230				235						240		
Asp	Leu	Asp	Pro	Gly	Gln	Thr	Leu	Ile	Tyr	Tyr	Val	Asp	Glu	Lys	Ala		
			245					250					255				
Pro	Glu	Phe	Ser	Met	Gln	Gly	Leu	Lys									
			260					265									

<210> 5

<211> 242

<212> PRT

<213> Artificial Sequence

<220>

<223> TAG-25 fragment comprising propeptide+ ECD

<400> 5

Gln	Glu	Glu	Cys	Val	Cys	Glu	Asn	Tyr	Lys	Leu	Ala	Val	Asn	Cys	Phe		
1				5					10					15			
Val	Asn	Asn	Asn	Arg	Glu	Cys	Gln	Cys	Thr	Ser	Val	Gly	Ala	Gln	Asn		
			20					25					30				
Thr	Val	Ile	Cys	Ser	Lys	Leu	Ala	Ala	Lys	Cys	Leu	Val	Met	Lys	Ala		
		35					40					45					
Glu	Met	Asn	Gly	Ser	Lys	Leu	Gly	Arg	Arg	Ile	Lys	Pro	Glu	Gly	Ala		
	50					55					60						
Leu	Gln	Asn	Asn	Asp	Gly	Leu	Tyr	Asp	Pro	Asp	Cys	Asp	Glu	Ser	Gly		
65				70				75							80		
Leu	Phe	Lys	Ala	Lys	Gln	Cys	Asn	Gly	Thr	Ala	Thr	Cys	Trp	Cys	Val		
			85					90					95				
Asn	Thr	Ala	Gly	Val	Arg	Arg	Thr	Asp	Lys	Asp	Thr	Glu	Ile	Thr	Cys		
		100						105					110				
Ser	Glu	Arg	Val	Arg	Thr	Tyr	Trp	Ile	Ile	Ile	Glu	Leu	Lys	His	Lys		
		115					120					125					
Glu	Arg	Glu	Ser	Pro	Tyr	Asp	Ser	Lys	Ser	Leu	His	Thr	Ala	Leu	Gln		
	130					135					140						
Lys	Glu	Ile	Thr	Thr	Arg	Tyr	Gln	Leu	Asp	Pro	Lys	Phe	Ile	Thr	Ser		
145					150				155						160		
Ile	Leu	Tyr	Glu	Asn	Asn	Val	Ile	Thr	Ile	Asp	Leu	Met	Gln	Asn	Ser		
			165					170					175				
Ser	Gln	Lys	Thr	Gln	Asp	Asp	Val	Asp	Ile	Ala	Asp	Val	Ala	Tyr	Tyr		

275 280 285
 Ser Arg Lys Lys Arg Met Ala Lys Tyr Glu Lys Ala Glu Ile Lys Glu
 290 295 300
 Met Gly Glu Met His Arg Glu Leu Asn Ala
 305 310

<210> 7
 <211> 234
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Mature Tag-25 polypeptide, which comprises N- to
 C-terminus ECD+TMD+CD

<220>
 <221> DOMAIN
 <222> (186)...(208)
 <223> Transmembrane domain (TMD)

<400> 7
 Arg Ile Lys Pro Glu Gly Ala Leu Gln Asn Asn Asp Gly Leu Tyr Asp
 1 5 10 15
 Pro Asp Cys Asp Glu Ser Gly Leu Phe Lys Ala Lys Gln Cys Asn Gly
 20 25 30
 Thr Ala Thr Cys Trp Cys Val Asn Thr Ala Gly Val Arg Arg Thr Asp
 35 40 45
 Lys Asp Thr Glu Ile Thr Cys Ser Glu Arg Val Arg Thr Tyr Trp Ile
 50 55 60
 Ile Ile Glu Leu Lys His Lys Glu Arg Glu Ser Pro Tyr Asp Ser Lys
 65 70 75 80
 Ser Leu His Thr Ala Leu Gln Lys Glu Ile Thr Thr Arg Tyr Gln Leu
 85 90 95
 Asp Pro Lys Phe Ile Thr Ser Ile Leu Tyr Glu Asn Asn Val Ile Thr
 100 105 110
 Ile Asp Leu Met Gln Asn Ser Ser Gln Lys Thr Gln Asp Asp Val Asp
 115 120 125
 Ile Ala Asp Val Ala Tyr Tyr Phe Glu Lys Asp Val Lys Gly Glu Ser
 130 135 140
 Leu Phe His Ser Lys Lys Met Asp Leu Arg Val Asn Gly Glu Leu Leu
 145 150 155 160
 Asp Leu Asp Pro Gly Gln Thr Leu Ile Tyr Tyr Val Asp Glu Lys Ala
 165 170 175
 Pro Glu Phe Ser Met Gln Gly Leu Lys Ala Gly Val Ile Ala Val Ile
 180 185 190
 Val Val Val Val Met Ala Val Val Ala Gly Ile Val Val Leu Val Ile
 195 200 205
 Ser Arg Lys Lys Arg Met Ala Lys Tyr Glu Lys Ala Glu Ile Lys Glu
 210 215 220
 Met Gly Glu Met His Arg Glu Leu Asn Ala
 225 230

<210> 8
 <211> 208
 <212> PRT

<213> Artificial Sequence

<220>

<223> Tag-25 fragment comprising ECD+TMD

<400> 8

```
Arg Ile Lys Pro Glu Gly Ala Leu Gln Asn Asn Asp Gly Leu Tyr Asp
 1           5           10           15
Pro Asp Cys Asp Glu Ser Gly Leu Phe Lys Ala Lys Gln Cys Asn Gly
          20           25           30
Thr Ala Thr Cys Trp Cys Val Asn Thr Ala Gly Val Arg Arg Thr Asp
          35           40           45
Lys Asp Thr Glu Ile Thr Cys Ser Glu Arg Val Arg Thr Tyr Trp Ile
 50           55           60
Ile Ile Glu Leu Lys His Lys Glu Arg Glu Ser Pro Tyr Asp Ser Lys
65           70           75           80
Ser Leu His Thr Ala Leu Gln Lys Glu Ile Thr Thr Arg Tyr Gln Leu
          85           90           95
Asp Pro Lys Phe Ile Thr Ser Ile Leu Tyr Glu Asn Asn Val Ile Thr
          100          105          110
Ile Asp Leu Met Gln Asn Ser Ser Gln Lys Thr Gln Asp Asp Val Asp
          115          120          125
Ile Ala Asp Val Ala Tyr Tyr Phe Glu Lys Asp Val Lys Gly Glu Ser
          130          135          140
Leu Phe His Ser Lys Lys Met Asp Leu Arg Val Asn Gly Glu Leu Leu
145          150          155          160
Asp Leu Asp Pro Gly Gln Thr Leu Ile Tyr Tyr Val Asp Glu Lys Ala
          165          170          175
Pro Glu Phe Ser Met Gln Gly Leu Lys Ala Gly Val Ile Ala Val Ile
          180          185          190
Val Val Val Val Met Ala Val Val Ala Gly Ile Val Val Leu Val Ile
          195          200          205
```

<210> 9

<211> 63

<212> PRT

<213> Artificial Sequence

<220>

<223> TAG-18 fragment comprising ECD

<400> 9

```
Gln Asn Asp Val Asp Ile Ala Asp Val Ala His Tyr Phe Glu Lys Asp
 1           5           10           15
Val Lys Gly Glu Ser Leu Phe His Ser Ser Lys Lys Met Asp Leu Arg
          20           25           30
Val Asn Gly Glu Gln Leu Asp Leu Asp Pro Gly Gln Thr Leu Ile Tyr
          35           40           45
Tyr Val Asp Arg Asn Ala Pro Glu Phe Ser Met Gln Ala Leu Lys
 50           55           60
```

<210> 10

<211> 115

<212> PRT

<213> Artificial Sequence

<220>

<223> TAg-18 ECD+TMD+CD

<400> 10

Gln Asn Asp Val Asp Ile Ala Asp Val Ala His Tyr Phe Glu Lys Asp
1 5 10 15
Val Lys Gly Glu Ser Leu Phe His Ser Ser Lys Lys Met Asp Leu Arg
20 25 30
Val Asn Gly Glu Gln Leu Asp Leu Asp Pro Gly Gln Thr Leu Ile Tyr
35 40 45
Tyr Val Asp Arg Asn Ala Pro Glu Phe Ser Met Gln Ala Leu Lys Ala
50 55 60
Gly Val Cys Ala Val Ile Val Val Val Met Ile Ala Val Val Ala Gly
65 70 75 80
Ile Val Val Leu Val Ile Ser Arg Lys Lys Arg Met Ala Lys Tyr Glu
85 90 95
Lys Ala Glu Ile Lys Glu Met Gly Arg Met His Arg Glu Leu Asn Ala
100 105 110
Ser Val Leu
115

<210> 11

<211> 29

<212> PRT

<213> Artificial Sequence

<220>

<223> TAg-18 CD-like sequence

<400> 11

Ser Arg Lys Lys Arg Met Ala Lys Tyr Glu Lys Ala Glu Ile Lys Glu
1 5 10 15
Met Gly Arg Met His Arg Glu Leu Asn Ala Ser Val Leu
20 25

<210> 12

<211> 186

<212> PRT

<213> Artificial Sequence

<220>

<223> TAg-21 fragment comprising ECD

<400> 12

Arg Ile Lys Pro Glu Gly Ala Leu Gln Asn Asn Asp Gly Leu Tyr Asp
1 5 10 15
Pro Asp Cys Asp Glu Ser Gly Leu Phe Lys Ala Lys Gln Cys Asn Gly
20 25 30
Thr Ala Thr Cys Trp Cys Val Asn Thr Ala Gly Val Arg Arg Thr Asp
35 40 45
Lys Asp Thr Glu Ile Thr Cys Ser Glu Arg Val Arg Thr Tyr Trp Ile
50 55 60
Ile Ile Glu Leu Lys His Lys Glu Arg Glu Ser Pro Tyr Asp Ser Lys
65 70 75 80

Ser	Leu	His	Thr	Ala	Leu	Gln	Lys	Glu	Ile	Thr	Thr	Arg	Tyr	Gln	Leu
				85					90					95	
Asp	Pro	Lys	Phe	Ile	Thr	Ser	Ile	Leu	Tyr	Glu	Asn	Asn	Val	Ile	Thr
			100					105					110		
Ile	Asp	Leu	Met	Gln	Asn	Ser	Ser	Gln	Lys	Thr	Gln	Asp	Asp	Val	Asp
		115					120					125			
Ile	Ala	Asp	Val	Ala	Tyr	Tyr	Phe	Glu	Lys	Asp	Val	Lys	Gly	Glu	Ser
	130						135				140				
Leu	Phe	His	Ser	Ser	Lys	Lys	Met	Asp	Leu	Arg	Val	Asn	Gly	Glu	Leu
145					150					155					160
Leu	Asp	Leu	Asp	Pro	Gly	Gln	Thr	Leu	Ile	Tyr	Tyr	Val	Asp	Glu	Lys
				165					170					175	
Ala	Pro	Glu	Phe	Ser	Met	Gln	Gly	Leu	Lys						
			180					185							

<210> 13

<211> 266

<212> PRT

<213> Artificial Sequence

<220>

<223> TAG-21 polypeptide (comprising SP+PP+ECD)

<400> 13

Met	Ala	Pro	Pro	Gln	Ala	Leu	Ala	Phe	Gly	Leu	Leu	Leu	Ala	Ala	Ala
1				5					10					15	
Thr	Ala	Thr	Phe	Ala	Ala	Ala	Gln	Glu	Glu	Cys	Val	Cys	Glu	Asn	Tyr
			20					25					30		
Lys	Leu	Ala	Val	Asn	Cys	Phe	Val	Asn	Asn	Asn	Arg	Glu	Cys	Gln	Cys
		35					40					45			
Thr	Ser	Val	Gly	Ala	Gln	Asn	Thr	Val	Ile	Cys	Ser	Lys	Leu	Ala	Ala
	50					55					60				
Lys	Cys	Leu	Val	Met	Lys	Ala	Glu	Met	Asn	Gly	Ser	Lys	Leu	Gly	Arg
65					70					75					80
Arg	Ile	Lys	Pro	Glu	Gly	Ala	Leu	Gln	Asn	Asn	Asp	Gly	Leu	Tyr	Asp
				85					90					95	
Pro	Asp	Cys	Asp	Glu	Ser	Gly	Leu	Phe	Lys	Ala	Lys	Gln	Cys	Asn	Gly
			100					105					110		
Thr	Ala	Thr	Cys	Trp	Cys	Val	Asn	Thr	Ala	Gly	Val	Arg	Arg	Thr	Asp
		115					120					125			
Lys	Asp	Thr	Glu	Ile	Thr	Cys	Ser	Glu	Arg	Val	Arg	Thr	Tyr	Trp	Ile
	130					135					140				
Ile	Ile	Glu	Leu	Lys	His	Lys	Glu	Arg	Glu	Ser	Pro	Tyr	Asp	Ser	Lys
145					150					155					160
Ser	Leu	Arg	Thr	Ala	Leu	Gln	Lys	Glu	Ile	Thr	Thr	Arg	Tyr	Gln	Leu
				165					170					175	
Asp	Pro	Lys	Phe	Ile	Thr	Ser	Ile	Leu	Tyr	Glu	Asn	Asn	Val	Ile	Thr
			180					185					190		
Ile	Asp	Leu	Met	Gln	Asn	Ser	Ser	Gln	Lys	Thr	Gln	Asn	Asp	Val	Asp
		195					200					205			
Ile	Ala	Asp	Val	Ala	His	Tyr	Phe	Glu	Lys	Asp	Val	Lys	Gly	Glu	Ser
	210					215					220				
Leu	Phe	His	Ser	Ser	Lys	Lys	Met	Asp	Leu	Arg	Val	Asn	Gly	Glu	Gln
225					230					235					240
Leu	Asp	Leu	Asp	Pro	Gly	Gln	Thr	Leu	Ile	Tyr	Tyr	Val	Asp	Arg	Asn
				245					250					255	

Ala Pro Glu Phe Ser Met Gln Ala Leu Lys
 260 265

<210> 14
 <211> 289
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> TAG-21 extended polypeptide comprising
 SP+PP+ECD+TMD

<400> 14

Met	Ala	Pro	Pro	Gln	Ala	Leu	Ala	Phe	Gly	Leu	Leu	Leu	Ala	Ala	Ala	1	5	10	15
Thr	Ala	Thr	Phe	Ala	Ala	Ala	Gln	Glu	Glu	Cys	Val	Cys	Glu	Asn	Tyr	20	25	30	
Lys	Leu	Ala	Val	Asn	Cys	Phe	Val	Asn	Asn	Asn	Arg	Glu	Cys	Gln	Cys	35	40	45	
Thr	Ser	Val	Gly	Ala	Gln	Asn	Thr	Val	Ile	Cys	Ser	Lys	Leu	Ala	Ala	50	55	60	
Lys	Cys	Leu	Val	Met	Lys	Ala	Glu	Met	Asn	Gly	Ser	Lys	Leu	Gly	Arg	65	70	75	80
Arg	Ile	Lys	Pro	Glu	Gly	Ala	Leu	Gln	Asn	Asn	Asp	Gly	Leu	Tyr	Asp	85	90	95	
Pro	Asp	Cys	Asp	Glu	Ser	Gly	Leu	Phe	Lys	Ala	Lys	Gln	Cys	Asn	Gly	100	105	110	
Thr	Ala	Thr	Cys	Trp	Cys	Val	Asn	Thr	Ala	Gly	Val	Arg	Arg	Thr	Asp	115	120	125	
Lys	Asp	Thr	Glu	Ile	Thr	Cys	Ser	Glu	Arg	Val	Arg	Thr	Tyr	Trp	Ile	130	135	140	
Ile	Ile	Glu	Leu	Lys	His	Lys	Glu	Arg	Glu	Ser	Pro	Tyr	Asp	Ser	Lys	145	150	155	160
Ser	Leu	Arg	Thr	Ala	Leu	Gln	Lys	Glu	Ile	Thr	Thr	Arg	Tyr	Gln	Leu	165	170	175	
Asp	Pro	Lys	Phe	Ile	Thr	Ser	Ile	Leu	Tyr	Glu	Asn	Asn	Val	Ile	Thr	180	185	190	
Ile	Asp	Leu	Met	Gln	Asn	Ser	Ser	Gln	Lys	Thr	Gln	Asn	Asp	Val	Asp	195	200	205	
Ile	Ala	Asp	Val	Ala	His	Tyr	Phe	Glu	Lys	Asp	Val	Lys	Gly	Glu	Ser	210	215	220	
Leu	Phe	His	Ser	Ser	Lys	Lys	Met	Asp	Leu	Arg	Val	Asn	Gly	Glu	Gln	225	230	235	240
Leu	Asp	Leu	Asp	Pro	Gly	Gln	Thr	Leu	Ile	Tyr	Tyr	Val	Asp	Arg	Asn	245	250	255	
Ala	Pro	Glu	Phe	Ser	Met	Gln	Ala	Leu	Lys	Ala	Gly	Ile	Ile	Ala	Val	260	265	270	
Ile	Val	Val	Val	Met	Ile	Ala	Val	Val	Ala	Gly	Ile	Val	Val	Leu	Val	275	280	285	
Ile																			

<210> 15
 <211> 23
 <212> PRT

<213> Artificial Sequence

<220>

<223> Tag-21 TMD

<400> 15

Ala Gly Ile Ile Ala Val Ile Val Val Val Met Ile Ala Val Val Ala
1 5 10 15
Gly Ile Val Val Leu Val Ile
20

<210> 16

<211> 555

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA encoding TAg-25 fragment comprising ECD domain

<400> 16

aggatcaaac ctgaaggagc tctgcagaac aacgatggtc tctacgaccc cgactgtgac 60
gagtcgggcc tcttcaaggc caaacagtgt aatggcactg ctacatgctg gtgcgtgaac 120
accgctgggg tgcgccggac cgataaggat accgaaatta cctgttctga gaggggccgg 180
acatattgga tcatcattga actcaaacat aaagagcgcg agtctccata cgattctaaa 240
tccctccata ctgcactgca aaaggaaatc actacacgct accagctgga tccaaaattc 300
attacatcca tcctctatga gaacaatgtt attacaattg atctgatgca aaatagctct 360
cagaagactc aagacgacgt ggacatcgct gatgtggcct actattttga gaaggacgtt 420
aagggggaat cactgttcca ttcaaagaaa atggatctga gggttaatgg cgagctgctg 480
gacctggacc cagggcaaac cctgatctat tatgtggacg agaaggctcc agaattctct 540
atgcaaggcc tgaag 555

<210> 17

<211> 171

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA encoding TAg-25 fragment comprising propeptide

<400> 17

caggaggagt gtgtgtgcca aaactacaag ctcgctgtca actgtttcgt caacaataat 60
agagaatgcc agtgcaattc tgtgggagca cagaatacag tgatctgtag caaactggct 120
gcaaagtgtc tggcatgaa ggccgaaatg aacggatcca agctcgggcg g 171

<210> 18

<211> 69

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA encoding TAg-25 fragment comprising SP

<400> 18

atggcacccc ctcaagcact ggcaactgggt ctgctgctgg ccgccgctac cgccactttc 60
gccgcagca 69

<210> 19
 <211> 795
 <212> DNA
 <213> Artificial Sequence

<220>

<223> DNA encoding TAg-25 polypeptide (which comprises
 signal peptide+propeptide+ ECD)

<400> 19

```
atggcacccc ctcaagcact ggcactgggt ctgctgctgg ccgccgctac cgccactttc 60
gccgcagcac aggaggagtg tgtgtgcgaa aactacaagc tcgctgtcaa ctgtttcgtc 120
aacaataata gagaatgcca gtgcacttct gtgggagcac agaatacagt gatctgtagc 180
aaactggctg caaagtgtct ggtcatgaag gccgaaatga acggatccaa gctcgggcgg 240
aggatcaaac ctgaaggagc tctgcagaac aacgatggtc tctacgacc cgactgtgac 300
gagtcgggcc tcttcaaggc caaacagtgt aatggcactg ctacatgctg gtgcgtgaac 360
accgctgggg tgcgccggac cgataaggat accgaaatta cctgttctga gaggggccgg 420
acatattgga tcatcattga actcaaacat aaagagcgcg agtctccata cgattctaaa 480
tccctccata ctgcactgca aaaggaaatc actacacgct accagctgga tccaaaattc 540
attacatcca tcctctatga gaacaatgtt attacaattg atctgatgca aaatagctct 600
cagaagactc aagacgacgt ggacatcgct gatgtggcct actattttga gaaggacgtt 660
aagggggaat cactgttcca ttcaaagaaa atggatctga ggggttaatg cgagctgctg 720
gacctggacc cagggcaaac cctgatctat tatgtggacg agaaggctcc agaattctct 780
atgcaaggcc tgaag 795
```

<210> 20
 <211> 726
 <212> DNA
 <213> Artificial Sequence

<220>

<223> DNA encoding TAg-25 fragment comprising PP+ECD

<400> 20

```
caggaggagt gtgtgtgcca aaactacaag ctcgctgtca actgtttcgt caacaataat 60
agagaatgcc agtgcacttc tgtgggagca cagaatacag tgatctgtag caaactggct 120
gcaaagtgtc tggatcatga ggccgaaatg aacggatcca agctcgggcg gaggatcaaa 180
cctgaaggag ctctgcagaa caacgatggc ctctacgacc ccgactgtga cgagtccggc 240
ctcttcaagg ccaaacagtg taatggcact gctacatgct ggtgcgtgaa caccgctggg 300
gtgcgccgga ccgataagga taccgaaatt acctgttctg agaggggccg gacatattgg 360
atcatcattg aactcaaaca taaagagcgc gagtctccat acgattctaa atccctccat 420
actgcactgc aaaaggaaat cactacacgc taccagctgg atccaaaatt cattacatcc 480
atcctctatg agaacaatgt tattacaatt gatctgatgc aaaatagctc tcagaagact 540
caagacgacg tggacatcgc tgatgtggcc tactattttg agaaggacgt taagggggaa 600
tcaactgttc attcaaagaa aatggatctg aggggttaat gcgagctgct ggacctggac 660
ccagggcaaa ccctgatcta ttatgtggac gagaaggctc cagaattctc tatgcaaggc 720
ctgaag 726
```

<210> 21
 <211> 942
 <212> DNA
 <213> Artificial Sequence

<220>

<223> DNA encoding TAg-25 full-length/membrane bound
 form which comprises SP+PP+ECD+TMD+CD

<220>
 <221> misc_feature
 <222> (796)...(864)
 <223> transmembrane domain (TMD) coding sequence

<220>
 <221> misc_feature
 <222> (865)...(942)
 <223> cytoplasmic domain (CD) coding sequence

<400> 21
 atggcacccc ctcaagcact ggcactgggt ctgctgctgg ccgcccgtac cgccactttc 60
 gccgcagcac aggaggagtg tgtgtgcgaa aactacaagc tcgctgtcaa ctgttttcgtc 120
 aacaataata gagaatgcca gtgcacttct gtgggagcac agaatacagt gatctgtagc 180
 aaactggctg caaagtgtct ggtcatgaag gccgaaatga acggatccaa gctcgggcgg 240
 aggatcaaac ctgaaggagc tctgcagaac aacgatggtc tctacgacc cgactgtgac 300
 gagtccggcc tcttcaaggc caaacagtgt aatggcactg ctacatgctg gtgcgtgaac 360
 accgctgggg tgccgcccgg cgataaggat accgaaatta cctgttctga gaggggtccg 420
 acatattgga tcatcattga actcaaacat aaagagcgcg agtctccata cgattctaaa 480
 tccctccata ctgcactgca aaaggaaatc actacacgct accagctgga tccaaaattc 540
 attacatcca tcctctatga gaacaatgtt attacaattg atctgatgca aaatagctct 600
 cagaagactc aagacgacgt ggacatcgct gatgtggcct actattttga gaaggacgtt 660
 aagggggaat cactgtttcca ttcaaagaaa atggatctga gggttaatgg cgagctgctg 720
 gacctggacc cagggcaaac cctgatctat tatgtggacg agaaggctcc agaattctct 780
 atgcaaggcc tgaaggctgg tgttattgct gttattgtgg ttgtggtgat ggcagttggt 840
 gctggaattg ttgtgctggg tatttccaga aagaagagaa tggcaaagta tgagaaggct 900
 gagataaagg agatgggtga gatgcatagg gaactcaatg ca 942

<210> 22
 <211> 702
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> DNA encoding TAg-25 Mature domain (which comprises ECD+TMD+CD of a full-length or membrane bound form of TAg-25).

<220>
 <221> misc_feature
 <222> (556)...(624)
 <223> transmembrane domain (TMD) coding sequence

<400> 22
 aggatcaaac ctgaaggagc tctgcagaac aacgatggtc tctacgacc cgactgtgac 60
 gagtccggcc tcttcaaggc caaacagtgt aatggcactg ctacatgctg gtgcgtgaac 120
 accgctgggg tgccgcccgg cgataaggat accgaaatta cctgttctga gaggggtccg 180
 acatattgga tcatcattga actcaaacat aaagagcgcg agtctccata cgattctaaa 240
 tccctccata ctgcactgca aaaggaaatc actacacgct accagctgga tccaaaattc 300
 attacatcca tcctctatga gaacaatgtt attacaattg atctgatgca aaatagctct 360
 cagaagactc aagacgacgt ggacatcgct gatgtggcct actattttga gaaggacgtt 420
 aagggggaat cactgtttcca ttcaaagaaa atggatctga gggttaatgg cgagctgctg 480
 gacctggacc cagggcaaac cctgatctat tatgtggacg agaaggctcc agaattctct 540
 atgcaaggcc tgaaggctgg tgttattgct gttattgtgg ttgtggtgat ggcagttggt 600
 gctggaattg ttgtgctggg tatttccaga aagaagagaa tggcaaagta tgagaaggct 660
 gagataaagg agatgggtga gatgcatagg gaactcaatg ca 702

<210> 23
 <211> 624
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> DNA encoding TAg-25 fragment comprising ECD+TMD

<220>
 <221> misc_feature
 <222> (556)...(624)
 <223> transmembrane domain (TMD) coding sequence

<400> 23
 aggatcaaac ctgaaggagc tctgcagaac aacgatggtc tctacgaccc cgactgtgac 60
 gagtccggcc tcttcaaggc caaacagtgt aatggcactg ctacatgctg gtgctgtaac 120
 accgctgggg tgcgccggac cgataaggat accgaaatta cctgttctga gaggggccgg 180
 acatattgga tcatcattga actcaaacat aaagagcgcg agtctccata cgattctaaa 240
 tccctccata ctgcactgca aaaggaaatc actacacgct accagctgga tccaaaattc 300
 attacatcca tcctctatga gaacaatggt attacaattg atctgatgca aaatagctct 360
 cagaagactc aagacgacgt ggacatcgct gatgtggcct actattttga gaaggacgtt 420
 aagggggaat cactgttcca ttcaaagaaa atggatctga gggttaatgg cgagctgctg 480
 gacctggacc cagggcaaac cctgatctat tatgtggacg agaaggctcc agaattctct 540
 atgcaaggcc tgaaggctgg tgttattgct gttattgtgg ttgtgggtgat ggcagttgtt 600
 gctggaattg ttgtgctggt tatt 624

<210> 24
 <211> 69
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> DNA sequence encoding TAg-21 fragment comprising
 SP

<400> 24
 atggcacctc cccaggcact ggcatttgga ctgctgctgg ctgcagcaac cgccacattc 60
 gctgctgcc 69

<210> 25
 <211> 171
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> DNA encoding TAg-21 fragment comprising propeptide

<400> 25
 caggaggagt gtgtgtgtga gaactataaa ctggctgtca attgttttgt taataacaat 60
 agggagtgcc aatgtactag cgtgggagcc caaaacactg tcatttgctc caaactcgcc 120
 gccaaatgct tcgtcatgaa agctgaaatg aatggtagca aactgggacg g 171

<210> 26
 <211> 558
 <212> DNA
 <213> Artificial Sequence

<220>

<223> DNA encoding Tag-21 fragment comprising ECD

<400> 26

```
aggattaagc ccgaaggggc cctccagaac aatgacggac tctacgatcc agactgcgac 60
gagagcgggc tgttcaaggc taagcagtgc aatggcaccg ccacctgttg gtgtgtgaat 120
accgctggag tgcggcggac agacaaagac actgagatca cctgtagcga gaggggtgcgc 180
acttattgga tcatcattga actgaaacac aaggaacgcg aatccccata tgattccaag 240
agcctgagga cgcgcctcca gaaagagatc actactagat atcagctgga ccccaaattc 300
atcaccagca ttctgtacga gaacaatgtc attacaatcg atctgatgca aaacagcagc 360
cagaagaccc agaatgacgt ggacatcgcc gatgtggccc attattttga gaaagatgtc 420
aagggggaat cactgttcca cagctccaag aagatggacc tgagagtga cgggtgaacaa 480
ctcgacctcg atcctgggca gacactgac tactatgtcg acaggaatgc ccctgaattc 540
agcatgcagg ccctgaag                                     558
```

<210> 27

<211> 867

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA encoding TAg-21 extended polypeptide
comprising SP+PP+ECD+TMD

<220>

<221> misc_feature

<222> (70)...(240)

<223> propeptide (PP) coding sequence

<220>

<221> misc_feature

<222> (799)...(867)

<223> transmembrane domain (TMD) coding sequence

<400> 27

```
atggcacctc cccaggcact ggcatttga ctgctgctgg ctgcagcaac cgccacattc 60
gctgctgccc aggaggagtg tgtgtgtgag aactataaac tggctgtcaa ttgttttgtt 120
aataacaata gggagtgcc aatgtactagc gtgggagccc aaaacactgt catttgctcc 180
aaactcgccg ccaaattgtc cgtcatgaaa gctgaaatga atggtagcaa actgggacgg 240
aggattaagc ccgaaggggc cctccagaac aatgacggac tctacgatcc agactgcgac 300
gagagcgggc tgttcaaggc taagcagtgc aatggcaccg ccacctgttg gtgtgtgaat 360
accgctggag tgcggcggac agacaaagac actgagatca cctgtagcga gaggggtgcgc 420
acttattgga tcatcattga actgaaacac aaggaacgcg aatccccata tgattccaag 480
agcctgagga cgcgcctcca gaaagagatc actactagat atcagctgga ccccaaattc 540
atcaccagca ttctgtacga gaacaatgtc attacaatcg atctgatgca aaacagcagc 600
cagaagaccc agaatgacgt ggacatcgcc gatgtggccc attattttga gaaagatgtc 660
aagggggaat cactgttcca cagctccaag aagatggacc tgagagtga cgggtgaacaa 720
ctcgacctcg atcctgggca gacactgac tactatgtcg acaggaatgc ccctgaattc 780
agcatgcagg ccctgaaggc cggatcatc gccgtgatcg tggttgttat gatcgccggt 840
tgggccggca tcgtcgtgct ggtgatc                                     867
```

<210> 28

<211> 345

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA sequence encoding TAg-18 alternative fragment
comprising ECD+TMD+CD

<400> 28

```
cagaatgatg tggacatagc tgatgtggct cattatcttg aaaaagatgt taaagggtgaa 60
tccttggttc attcttctaa gaaaatggac ctgagagtaa atggagaaca actggatctg 120
gatcctggtc aaactttaat ttattatggt gatagaaatg cacctgaatt ttcaatgcag 180
gctctaaaag ctggtgtttg tgctgttatt gtggttggtga tgatagcagt tgttgctgga 240
attgttgtgc tggttatttc cagaaagaag agaatggcaa agtatgagaa ggctgagata 300
aaggagatgg gtaggatgca tagggaactc aatgcatcag tccta 345
```

<210> 29

<211> 552

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA sequence encoding human EpCAM fragment
comprising hEpCAM ECD

<400> 29

```
gcaaaacctg aaggggccct ccagaacaat gatgggcttt atgacccctga ctgcgatgag 60
agcgggctct ttaaggccaa gcagtgaac ggcacctcca cgtgctgggtg tgtgaacact 120
gctgggggtca gaagaacaga caaggacact gaaataacct gctctgagcg agtgagaacc 180
tactggatca tcattgaact aaaacacaaa gcaagagaaa aaccttatga tagtaaaagt 240
ttgcggactg cacttcagaa ggagatcaca acgcgttatc aactggatcc aaaatttatc 300
acgagtatct tgtatgagaa taatgttatc actattgatc tggttcaaaa ttcttctcaa 360
aaaactcaga atgatgtgga catagctgat gtggcttatt attttgaaaa agatgttaaa 420
ggtgaatcct tgtttcattc taagaaaatg gacctgacag taaatgggga acaactggat 480
ctggatcctg gtcaaaacttt aatttattat gttgatgaaa aagcacctga attctcaatg 540
cagggtctaa aa 552
```

<210> 30

<211> 63

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA encoding hEpCAM fragment comprising signal
peptide

<400> 30

```
atggcgcccc cgcaggctct cgcgttcggg cttctgcttg ccgcggcgac ggcgactttt 60
gcc 63
```

<210> 31

<211> 180

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA encoding hEpCAM fragment comprising propeptide

<400> 31

```
gcagctcagg aagaatgtgt ctgtgaaaac tacaagctgg ccgtaaactg ctttgtgaat 60
aataatcgtc aatgccagtg tacttcagtt ggtgcacaaa atactgtcat ttgctcaaa 120
```

ctggctgccca aatgttttgggt gatgaaggca gaaatgaatg gctcaaaaact tgggagaaga 180

<210> 32

<211> 266

<212> PRT

<213> Artificial Sequence

<220>

<223> TAg-18 chimeric polypeptide (comprising SP+PP+ECD)

<400> 32

```

Met Ala Pro Pro Gln Val Leu Ala Phe Gly Leu Leu Leu Ala Ala Ala
 1              5              10              15
Thr Ala Thr Phe Ala Ala Ala Gln Glu Cys Val Cys Glu Asn Tyr
      20              25              30
Lys Leu Ala Val Asn Cys Phe Val Asn Asn Asn Arg Gln Cys Gln Cys
      35              40              45
Thr Ser Val Gly Ala Gln Asn Thr Val Ile Cys Ser Lys Leu Ala Ala
      50              55              60
Lys Cys Leu Val Met Lys Ala Glu Met Asn Gly Ser Lys Leu Gly Arg
65              70              75              80
Arg Ala Lys Pro Glu Gly Ala Leu Gln Asn Asn Asp Gly Leu Tyr Asp
      85              90              95
Pro Asp Cys Asp Glu Ser Gly Leu Phe Lys Ala Lys Gln Cys Asn Gly
      100             105             110
Thr Ser Thr Cys Trp Cys Val Asn Thr Ala Gly Val Arg Arg Thr Asp
      115             120             125
Lys Asp Thr Glu Ile Thr Cys Ser Glu Arg Val Arg Thr Tyr Trp Ile
      130             135             140
Ile Ile Glu Leu Lys His Lys Ala Arg Glu Lys Pro Tyr Asp Ser Lys
145             150             155             160
Ser Leu Arg Thr Ala Leu Gln Lys Glu Ile Thr Thr Arg Tyr Gln Leu
      165             170             175
Asp Pro Lys Phe Ile Thr Ser Ile Leu Tyr Glu Asn Asn Val Ile Thr
      180             185             190
Ile Asp Leu Val Gln Asn Ser Ser Gln Lys Thr Gln Asn Asp Val Asp
      195             200             205
Ile Ala Asp Val Ala His Tyr Phe Glu Lys Asp Val Lys Gly Glu Ser
      210             215             220
Leu Phe His Ser Ser Lys Lys Met Asp Leu Arg Val Asn Gly Glu Gln
225             230             235             240
Leu Asp Leu Asp Pro Gly Gln Thr Leu Ile Tyr Tyr Val Asp Arg Asn
      245             250             255
Ala Pro Glu Phe Ser Met Gln Ala Leu Lys
      260             265

```

<210> 33

<211> 798

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA encoding TAg-18 chimeric polypeptide
comprising SP+PP+ECD

<400> 33

```
atggcgcccc cgcaggtcct cgcgttcggg cttctgcttg ccgcggcgac ggcgactttt 60
gccgcagctc aggaagaatg tgtctgtgaa aactacaagc tggccgtaaa ctgctttgtg 120
aataataatc gtcaatgcc a gtgtacttca gttggtgcac aaaatactgt catttgctca 180
aagctggctg ccaaagtgtt ggtgatgaag gcagaaatga atggctcaaa acttgggaga 240
agagcaaaac ctgaaggggc cctccagaac aatgatgggc tttatgatcc tgactgcgat 300
gagagcgggc tctttaaggc caagcagtgc aacggcacct ccacgtgctg gtgtgtgaac 360
actgctgggg tcagaagaac agacaaggac actgaaataa cctgctctga gcgagtga 420
acctactgga tcatcattga actaaaacac aaagcaagag aaaaacctta tgatagtaaa 480
agtttgcgga ctgcacttca gaaggagatc acaacgcgtt atcaactgga tccaaaattt 540
atcacgagta ttttgtatga gaataatgtt atcactattg atctggttca aaattcttct 600
caaaaaactc agaatgatgt ggacatagct gatgtggctc attattttga aaaagatgtt 660
aaaggtgaat ccttgtttca tttctctaag aaaatggacc tgagagtaaa tggagaacaa 720
ctggatctgg atcctggta aactttaatt tattatgttg atagaaatgc acctgaattt 780
tcaatgcagg ctctaaaa
```

<210> 34

<211> 318

<212> PRT

<213> Artificial Sequence

<220>

<223> TAg-18 chimera extended polypeptide which
comprises SP+PP+ECD+TMD+CD

<400> 34

```
Met Ala Pro Pro Gln Val Leu Ala Phe Gly Leu Leu Leu Ala Ala Ala
1      5      10      15
Thr Ala Thr Phe Ala Ala Ala Gln Glu Cys Val Cys Glu Asn Tyr
20     25     30
Lys Leu Ala Val Asn Cys Phe Val Asn Asn Asn Arg Gln Cys Gln Cys
35     40     45
Thr Ser Val Gly Ala Gln Asn Thr Val Ile Cys Ser Lys Leu Ala Ala
50     55     60
Lys Cys Leu Val Met Lys Ala Glu Met Asn Gly Ser Lys Leu Gly Arg
65     70     75     80
Arg Ala Lys Pro Glu Gly Ala Leu Gln Asn Asn Asp Gly Leu Tyr Asp
85     90     95
Pro Asp Cys Asp Glu Ser Gly Leu Phe Lys Ala Lys Gln Cys Asn Gly
100    105    110
Thr Ser Thr Cys Trp Cys Val Asn Thr Ala Gly Val Arg Arg Thr Asp
115    120    125
Lys Asp Thr Glu Ile Thr Cys Ser Glu Arg Val Arg Thr Tyr Trp Ile
130    135    140
Ile Ile Glu Leu Lys His Lys Ala Arg Glu Lys Pro Tyr Asp Ser Lys
145    150    155    160
Ser Leu Arg Thr Ala Leu Gln Lys Glu Ile Thr Thr Arg Tyr Gln Leu
165    170    175
Asp Pro Lys Phe Ile Thr Ser Ile Leu Tyr Glu Asn Asn Val Ile Thr
180    185    190
Ile Asp Leu Val Gln Asn Ser Ser Gln Lys Thr Gln Asn Asp Val Asp
195    200    205
Ile Ala Asp Val Ala His Tyr Phe Glu Lys Asp Val Lys Gly Glu Ser
210    215    220
Leu Phe His Ser Ser Lys Lys Met Asp Leu Arg Val Asn Gly Glu Gln
225    230    235    240
Leu Asp Leu Asp Pro Gly Gln Thr Leu Ile Tyr Tyr Val Asp Arg Asn
```


65					70					75					80
Ser	Leu	Arg	Thr	Ala	Leu	Gln	Lys	Glu	Ile	Thr	Thr	Arg	Tyr	Gln	Leu
				85					90					95	
Asp	Pro	Lys	Phe	Ile	Thr	Ser	Ile	Leu	Tyr	Glu	Asn	Asn	Val	Ile	Thr
			100					105					110		
Ile	Asp	Leu	Val	Gln	Asn	Ser	Ser	Gln	Lys	Thr	Gln	Asn	Asp	Val	Asp
		115					120					125			
Ile	Ala	Asp	Val	Ala	Tyr	Tyr	Phe	Glu	Lys	Asp	Val	Lys	Gly	Glu	Ser
	130						135				140				
Leu	Phe	His	Ser	Lys	Lys	Met	Asp	Leu	Thr	Val	Asn	Gly	Glu	Gln	Leu
145					150					155					160
Asp	Leu	Asp	Pro	Gly	Gln	Thr	Leu	Ile	Tyr	Tyr	Val	Asp	Glu	Lys	Ala
				165					170					175	
Pro	Glu	Phe	Ser	Met	Gln	Gly	Leu	Lys							
			180					185							

<210> 37
 <211> 23
 <212> PRT
 <213> Homo sapiens

<220>
 <223> Fragment of hEpCAM comprising hEpCAM signal peptide

<400> 37
Met Ala Pro Pro Gln Val Leu Ala Phe Gly Leu Leu Leu Ala Ala Ala
1 5 10 15
Thr Ala Thr Phe Ala Ala Ala
20

<210> 38
 <211> 57
 <212> PRT
 <213> Homo sapiens

<220>
 <223> Fragment of hEpCAM comprising hEpCAM propeptide

<400> 38
Gln Glu Glu Cys Val Cys Glu Asn Tyr Lys Leu Ala Val Asn Cys Phe
1 5 10 15
Val Asn Asn Asn Arg Gln Cys Gln Cys Thr Ser Val Gly Ala Gln Asn
20 25 30
Thr Val Ile Cys Ser Lys Leu Ala Ala Lys Cys Leu Val Met Lys Ala
35 40 45
Glu Met Asn Gly Ser Lys Leu Gly Arg
50 55

<210> 39
 <211> 242
 <212> PRT
 <213> Homo sapiens

<220>

<223> Fragment of hEpCAM comprising hEpCAM propeptide+
ECD

<400> 39

Gln Glu Glu Cys Val Cys Glu Asn Tyr Lys Leu Ala Val Asn Cys Phe
1 5 10 15
Val Asn Asn Asn Arg Gln Cys Gln Cys Thr Ser Val Gly Ala Gln Asn
20 25 30
Thr Val Ile Cys Ser Lys Leu Ala Ala Lys Cys Leu Val Met Lys Ala
35 40 45
Glu Met Asn Gly Ser Lys Leu Gly Arg Arg Ala Lys Pro Glu Gly Ala
50 55 60
Leu Gln Asn Asn Asp Gly Leu Tyr Asp Pro Asp Cys Asp Glu Ser Gly
65 70 75 80
Leu Phe Lys Ala Lys Gln Cys Asn Gly Thr Ser Thr Cys Trp Cys Val
85 90 95
Asn Thr Ala Gly Val Arg Arg Thr Asp Lys Asp Thr Glu Ile Thr Cys
100 105 110
Ser Glu Arg Val Arg Thr Tyr Trp Ile Ile Ile Glu Leu Lys His Lys
115 120 125
Ala Arg Glu Lys Pro Tyr Asp Ser Lys Ser Leu Arg Thr Ala Leu Gln
130 135 140
Lys Glu Ile Thr Thr Arg Tyr Gln Leu Asp Pro Lys Phe Ile Thr Ser
145 150 155 160
Ile Leu Tyr Glu Asn Asn Val Ile Thr Ile Asp Leu Val Gln Asn Ser
165 170 175
Ser Gln Lys Thr Gln Asn Asp Val Asp Ile Ala Asp Val Ala Tyr Tyr
180 185 190
Phe Glu Lys Asp Val Lys Gly Glu Ser Leu Phe His Ser Lys Lys Met
195 200 205
Asp Leu Thr Val Asn Gly Glu Gln Leu Asp Leu Asp Pro Gly Gln Thr
210 215 220
Leu Ile Tyr Tyr Val Asp Glu Lys Ala Pro Glu Phe Ser Met Gln Gly
225 230 235 240
Leu Lys

<210> 40

<211> 265

<212> PRT

<213> Homo sapiens

<220>

<223> Fragment of hEpCAM comprising hEpCAM signal
peptide+ propeptide+ECD (termed "sEpCAM")

<400> 40

Met Ala Pro Pro Gln Val Leu Ala Phe Gly Leu Leu Leu Ala Ala Ala
1 5 10 15
Thr Ala Thr Phe Ala Ala Ala Gln Glu Glu Cys Val Cys Glu Asn Tyr
20 25 30
Lys Leu Ala Val Asn Cys Phe Val Asn Asn Asn Arg Gln Cys Gln Cys
35 40 45
Thr Ser Val Gly Ala Gln Asn Thr Val Ile Cys Ser Lys Leu Ala Ala
50 55 60

Lys	Cys	Leu	Val	Met	Lys	Ala	Glu	Met	Asn	Gly	Ser	Lys	Leu	Gly	Arg	65	70	75	80
Arg	Ala	Lys	Pro	Glu	Gly	Ala	Leu	Gln	Asn	Asn	Asp	Gly	Leu	Tyr	Asp	85	90	95	
Pro	Asp	Cys	Asp	Glu	Ser	Gly	Leu	Phe	Lys	Ala	Lys	Gln	Cys	Asn	Gly	100	105	110	
Thr	Ser	Thr	Cys	Trp	Cys	Val	Asn	Thr	Ala	Gly	Val	Arg	Arg	Thr	Asp	115	120	125	
Lys	Asp	Thr	Glu	Ile	Thr	Cys	Ser	Glu	Arg	Val	Arg	Thr	Tyr	Trp	Ile	130	135	140	
Ile	Ile	Glu	Leu	Lys	His	Lys	Ala	Arg	Glu	Lys	Pro	Tyr	Asp	Ser	Lys	145	150	155	160
Ser	Leu	Arg	Thr	Ala	Leu	Gln	Lys	Glu	Ile	Thr	Thr	Arg	Tyr	Gln	Leu	165	170	175	
Asp	Pro	Lys	Phe	Ile	Thr	Ser	Ile	Leu	Tyr	Glu	Asn	Asn	Val	Ile	Thr	180	185	190	
Ile	Asp	Leu	Val	Gln	Asn	Ser	Ser	Gln	Lys	Thr	Gln	Asn	Asp	Val	Asp	195	200	205	
Ile	Ala	Asp	Val	Ala	Tyr	Tyr	Phe	Glu	Lys	Asp	Val	Lys	Gly	Glu	Ser	210	215	220	
Leu	Phe	His	Ser	Lys	Lys	Met	Asp	Leu	Thr	Val	Asn	Gly	Glu	Gln	Leu	225	230	235	240
Asp	Leu	Asp	Pro	Gly	Gln	Thr	Leu	Ile	Tyr	Tyr	Val	Asp	Glu	Lys	Ala	245	250	255	
Pro	Glu	Phe	Ser	Met	Gln	Gly	Leu	Lys								260	265		

<210> 41

<211> 314

<212> PRT

<213> Homo sapiens

<220>

<223> WT full-length / membrane-bound hEpCAM (comprising
SP+PP+ECD+TMD+CD domains)

<220>

<221> DOMAIN

<222> (1)...(23)

<223> signal peptide (SP)

<220>

<221> DOMAIN

<222> (81)...(265)

<223> extracellular domain (ECD)

<220>

<221> DOMAIN

<222> (266)...(288)

<223> transmembrane domain (TMD)

<220>

<221> DOMAIN

<222> (289)...(314)

<223> cytoplasmic domain (CD)

<400> 41

```
Met Ala Pro Pro Gln Val Leu Ala Phe Gly Leu Leu Leu Ala Ala Ala
 1          5          10          15
Thr Ala Thr Phe Ala Ala Ala Gln Glu Glu Cys Val Cys Glu Asn Tyr
      20          25          30
Lys Leu Ala Val Asn Cys Phe Val Asn Asn Asn Arg Gln Cys Gln Cys
      35          40          45
Thr Ser Val Gly Ala Gln Asn Thr Val Ile Cys Ser Lys Leu Ala Ala
      50          55          60
Lys Cys Leu Val Met Lys Ala Glu Met Asn Gly Ser Lys Leu Gly Arg
65          70          75          80
Arg Ala Lys Pro Glu Gly Ala Leu Gln Asn Asn Asp Gly Leu Tyr Asp
      85          90          95
Pro Asp Cys Asp Glu Ser Gly Leu Phe Lys Ala Lys Gln Cys Asn Gly
      100          105          110
Thr Ser Thr Cys Trp Cys Val Asn Thr Ala Gly Val Arg Arg Thr Asp
      115          120          125
Lys Asp Thr Glu Ile Thr Cys Ser Glu Arg Val Arg Thr Tyr Trp Ile
      130          135          140
Ile Ile Glu Leu Lys His Lys Ala Arg Glu Lys Pro Tyr Asp Ser Lys
145          150          155          160
Ser Leu Arg Thr Ala Leu Gln Lys Glu Ile Thr Thr Arg Tyr Gln Leu
      165          170          175
Asp Pro Lys Phe Ile Thr Ser Ile Leu Tyr Glu Asn Asn Val Ile Thr
      180          185          190
Ile Asp Leu Val Gln Asn Ser Ser Gln Lys Thr Gln Asn Asp Val Asp
      195          200          205
Ile Ala Asp Val Ala Tyr Tyr Phe Glu Lys Asp Val Lys Gly Glu Ser
      210          215          220
Leu Phe His Ser Lys Lys Met Asp Leu Thr Val Asn Gly Glu Gln Leu
225          230          235          240
Asp Leu Asp Pro Gly Gln Thr Leu Ile Tyr Tyr Val Asp Glu Lys Ala
      245          250          255
Pro Glu Phe Ser Met Gln Gly Leu Lys Ala Gly Val Ile Ala Val Ile
      260          265          270
Val Val Val Val Met Ala Val Val Ala Gly Ile Val Val Leu Val Ile
      275          280          285
Ser Arg Lys Lys Arg Met Ala Lys Tyr Glu Lys Ala Glu Ile Lys Glu
      290          295          300
Met Gly Glu Met His Arg Glu Leu Asn Ala
305          310
```

<210> 42

<211> 942

<212> DNA

<213> Homo sapiens

<220>

<223> DNA encoding WT full-length/membrane-bound hEpCAM
(comprising signal peptide+propeptide+ECD+TMD+CD
domains)

<220>

<221> misc_feature

<222> (1)...(63)

<223> coding sequence for SP

<220>
 <221> misc_feature
 <222> (244)...(795)
 <223> coding sequence for ECD

<220>
 <221> misc_feature
 <222> (865)...(942)
 <223> coding sequence for CD

<400> 42
 atggcgcccc cgcaggtcct cgcgttcggg cttctgcttg ccgcggcgac ggcgactttt 60
 gccgcagctc aggaagaatg tgtctgtgaa aactacaagc tggccgtaaa ctgctttgtg 120
 aataataatc gtcaatgccca gtgtacttca gttggtgcac aaaataactgt catttgcctca 180
 aagctggctg ccaaagtgtt ggtgatgaag gcagaaatga atggctcaaa acttggggaga 240
 agagcaaaac ctgaaggggc cctccagaac aatgatgggc tttatgatcc tgactgcgat 300
 gagagcgggc tctttaaggc caagcagtcg aacggcacct ccacgtgctg gtgtgtgaac 360
 actgctgggg tcagaagaac agacaaggac actgaaataa cctgctctga gcgagtgaga 420
 acctactgga tcatcattga actaaaacac aaagcaagag aaaaacctta tgatagtaaa 480
 agtttgcgga ctgcacttca gaaggagatc acaacgcgtt atcaactgga tccaaaattt 540
 atcacgagta ttttgtatga gaataatgtt atcactattg atctggttca aaattcttct 600
 caaaaaactc agaatgatgt ggacatagct gatgtggctt attattttga aaaagatgtt 660
 aaaggtgaat ccttgtttca ttctaagaaa atggacctga cagtaaatgg ggaacaactg 720
 gatctggatc ctgggtcaaac tttaatttat tatgttgatg aaaaagcacc tgaattctca 780
 atgcagggtc taaaagctgg tgttattgct gttattgtgg ttgtgggtgat ggcagttgtt 840
 gctggaattg ttgtgctggg tatttccaga aagaagagaa tggcaaagta tgagaaggct 900
 gagataaagg agatgggtga gatgcatagg gaactcaatg ca 942

<210> 43
 <211> 234
 <212> PRT
 <213> Homo sapiens

<220>
 <223> Mature domain of hEpCAM (comprising ECD+TMD+CD)

<220>
 <221> DOMAIN
 <222> (186)...(208)
 <223> transmembrane domain (TMD)

<400> 43
 Arg Ala Lys Pro Glu Gly Ala Leu Gln Asn Asn Asp Gly Leu Tyr Asp
 1 5 10 15
 Pro Asp Cys Asp Glu Ser Gly Leu Phe Lys Ala Lys Gln Cys Asn Gly
 20 25 30
 Thr Ser Thr Cys Trp Cys Val Asn Thr Ala Gly Val Arg Arg Thr Asp
 35 40 45
 Lys Asp Thr Glu Ile Thr Cys Ser Glu Arg Val Arg Thr Tyr Trp Ile
 50 55 60
 Ile Ile Glu Leu Lys His Lys Ala Arg Glu Lys Pro Tyr Asp Ser Lys
 65 70 75 80
 Ser Leu Arg Thr Ala Leu Gln Lys Glu Ile Thr Thr Arg Tyr Gln Leu
 85 90 95
 Asp Pro Lys Phe Ile Thr Ser Ile Leu Tyr Glu Asn Asn Val Ile Thr
 100 105 110

Ile Asp Leu Val Gln Asn Ser Ser Gln Lys Thr Gln Asn Asp Val Asp
115 120 125
Ile Ala Asp Val Ala Tyr Tyr Phe Glu Lys Asp Val Lys Gly Glu Ser
130 135 140
Leu Phe His Ser Lys Lys Met Asp Leu Thr Val Asn Gly Glu Gln Leu
145 150 155 160
Asp Leu Asp Pro Gly Gln Thr Leu Ile Tyr Tyr Val Asp Glu Lys Ala
165 170 175
Pro Glu Phe Ser Met Gln Gly Leu Lys Ala Gly Val Ile Ala Val Ile
180 185 190
Val Val Val Val Met Ala Val Val Ala Gly Ile Val Val Leu Val Ile
195 200 205
Ser Arg Lys Lys Arg Met Ala Lys Tyr Glu Lys Ala Glu Ile Lys Glu
210 215 220
Met Gly Glu Met His Arg Glu Leu Asn Ala
225 230

<210> 44
<211> 699
<212> DNA
<213> Homo sapiens

<220>
<223> DNA encoding mature domain of hEpCAM (comprising
ECD+TMD+CD)

<220>
<221> misc_feature
<222> (553)...(621)
<223> coding sequence for TMD

<400> 44
gcaaaacctg aaggggccct ccagaacaat gatgggcttt atgatcctga ctgcatgag 60
agcgggctct ttaaggccaa gcagtgaac ggcacctcca cgtgctgggtg tgtgaacact 120
gctgggggtca gaagaacaga caaggacact gaaataacct gctctgagcg agtgagaacc 180
tactggatca tcattgaact aaaacacaaa gcaagagaaa aaccttatga tagtaaaagt 240
ttgcggactg cacttcagaa ggagatcaca acgcgttatc aactggatcc aaaattttatc 300
acgagtattt tgtatgagaa taatgttatc actattgatc tggttcaaaa ttcttctcaa 360
aaaactcaga atgatgtgga catagctgat gtggcttatt attttgaaaa agatgttaaa 420
ggtgaatcct tgtttcattc taagaaaatg gacctgacag taaatgggga acaactggat 480
ctggatcctg gtcaaaacttt aatttattat gttgatgaaa aagcacctga attctcaatg 540
caggggtctaa aagctgggtg tattgctgtt attgtgggtg tggatgatggc agttgttgct 600
ggaattgttg tgctgggttat ttccagaaag aagagaatgg caaagtatga gaaggctgag 660
ataaaggaga tgggtgagat gcatagggaa ctcaatgca 699

<210> 45
<211> 23
<212> PRT
<213> Artificial Sequence

<220>
<223> Fragment of hEpCAM comprising TMD

<400> 45
Ala Gly Val Ile Ala Val Ile Val Val Val Val Met Ala Val Val Ala
1 5 10 15

Gly Ile Val Val Leu Val Ile
20

<210> 46
<211> 26
<212> PRT
<213> Artificial Sequence

<220>
<223> Fragment of hEpCAM CD

<400> 46
Ser Arg Lys Lys Arg Met Ala Lys Tyr Glu Lys Ala Glu Ile Lys Glu
1 5 10 15
Met Gly Glu Met His Arg Glu Leu Asn Ala
20 25

<210> 47
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope of hEpCAM in ECD

<400> 47
Gly Leu Tyr Asp Pro Asp Cys Asp Glu
1 5

<210> 48
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope of hEpCAM in ECD

<400> 48
Ile Leu Tyr Glu Asn Asn Val Ile Thr
1 5

<210> 49
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope of hEpCAM in ECD

<400> 49
Tyr Gln Leu Asp Pro Lys Phe Ile Thr
1 5

<210> 50
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope of hEpCAM in ECD

<400> 50
Tyr Gln Leu Asp Pro Lys Phe Ile Thr Ser Ile
1 5 10

<210> 51
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope of hEpCAM in ECD

<400> 51
Tyr Gln Leu Asp Pro Lys Phe Ile Thr Ser Ile Leu Tyr Glu Asn Asn
1 5 10 15
Val Ile Thr

<210> 52
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope of hEpCAM in ECD

<400> 52
Tyr Gln Leu Asp Pro Lys Phe Ile Thr Ser Ile Leu Tyr Glu Asn Asn
1 5 10 15
Val Ile Thr Ile
20

<210> 53
<211> 21
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope - in ECD

<400> 53
Tyr Gln Leu Asp Pro Lys Phe Ile Thr Ser Ile Leu Tyr Glu Asn Asn
1 5 10 15
Val Ile Thr Ser Ile
20

<210> 54
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope - in ECD

<400> 54
Leu Asp Leu Asp Pro Gly Gln Thr Leu
1 5

<210> 55
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope - in ECD

<400> 55
Leu Leu Asp Leu Asp Pro Gly Gln Thr Leu
1 5 10

<210> 56
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope - in ECD

<400> 56
Gln Leu Asp Leu Asp Pro Gly Gln Thr Leu
1 5 10

<210> 57
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope - in ECD

<400> 57
Trp Ile Ile Ile Glu Leu Lys His Lys Ala
1 5 10

<210> 58
<211> 10
<212> PRT

<213> Artificial Sequence

<220>

<223> Epitope - in ECD

<400> 58

Trp Ile Ile Ile Glu Leu Lys His Lys Glu
1 5 10

<210> 59

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Epitope - in ECD

<400> 59

Ser Thr Cys Trp Cys Val Asn Thr Ala
1 5

<210> 60

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Epitope - in ECD

<400> 60

Ala Thr Cys Trp Cys Val Asn Thr Ala
1 5

<210> 61

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> Epitope - in ECD

<400> 61

Tyr Val Asp Glu Lys Ala Pro Glu Phe Ser Met
1 5 10

<210> 62

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> Epitope - in ECD

<400> 62

Tyr Val Asp Glu Lys Ala Pro Glu Phe Ser Asn
1 5 10

<210> 63

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> Epitope - in ECD

<400> 63

Gln Asn Asn Asp Gly Leu Tyr Asp Pro Asp Cys Asp Glu Ser Gly Leu
1 5 10 15
Phe Asp

<210> 64

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> Epitope - in ECD

<400> 64

Gln Asn Asn Asp Gly Leu Tyr Asp Pro Asp Cys Asp Glu Ser Gly Leu
1 5 10 15
Phe Lys

<210> 65

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Epitope in ECD/TMD

<400> 65

Ser Met Gln Gly Leu Lys Ala Gly Val
1 5

<210> 66

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> Epitope in ECD/TMD

<400> 66

Ser Met Gln Gly Leu Lys Ala Val Ala Gly Val
1 5 10

<210> 67
<211> 14
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope in ECD/TMD

<400> 67
Ser Met Gln Gly Leu Lys Ala Val Ala Gly Val Thr Ala Val
1 5 10

<210> 68
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope in ECD/TMD

<400> 68
Gly Leu Lys Ala Gly Val Ile Ala Val Ile Val
1 5 10

<210> 69
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope in ECD/TMD

<400> 69
Gly Leu Lys Ala Gly Val Ile Ala Val
1 5

<210> 70
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope in ECD/TMD

<400> 70
Gly Leu Lys Ala Gly Val Ile Ala Val Ile
1 5 10

<210> 71

<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Propeptide epitope

<400> 71
Cys Val Cys Glu Asn Tyr Lys Leu Ala Val
1 5 10

<210> 72
<211> 18
<212> PRT
<213> Artificial Sequence

<220>
<223> Propeptide epitope

<400> 72
Ala Gln Asn Thr Val Ile Cys Ser Lys Leu Ala Ala Lys Cys Leu Val
1 5 10 15
Met Lys

<210> 73
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Propeptide epitope

<400> 73
Leu Leu Leu Ala Ala Ala Thr Ala Thr Phe Ala
1 5 10

<210> 74
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Epitope - in signal peptide

<400> 74
Gln Val Leu Ala Phe Gly Leu Leu Leu
1 5

<210> 75
<211> 9
<212> PRT
<213> Artificial Sequence

<220>

<223> Epitope - in signal peptide

<400> 75

Leu Leu Ala Ala Thr Ala Thr Phe Ala
1 5

<210> 76

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> TAg-25 signal peptide epitope

<400> 76

Gln Ala Leu Ala Leu Gly Leu Leu Leu
1 5

<210> 77

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Epitope in TMD

<400> 77

Val Val Ala Gly Ile Val Val Leu Val
1 5

<210> 78

<211> 266

<212> PRT

<213> Artificial Sequence

<220>

<223> Tag-25/18 chimera (comprising SP+PP+ECD)

<400> 78

Met Ala Pro Pro Gln Ala Leu Ala Leu Gly Leu Leu Leu Ala Ala Ala
1 5 10 15
Thr Ala Thr Phe Ala Ala Ala Gln Glu Glu Cys Val Cys Glu Asn Tyr
20 25 30
Lys Leu Ala Val Asn Cys Phe Val Asn Asn Asn Arg Glu Cys Gln Cys
35 40 45
Thr Ser Val Gly Ala Gln Asn Thr Val Ile Cys Ser Lys Leu Ala Ala
50 55 60
Lys Cys Leu Val Met Lys Ala Glu Met Asn Gly Ser Lys Leu Gly Arg
65 70 75 80
Arg Ile Lys Pro Glu Gly Ala Leu Gln Asn Asn Asp Gly Leu Tyr Asp
85 90 95
Pro Asp Cys Asp Glu Ser Gly Leu Phe Lys Ala Lys Gln Cys Asn Gly

	100		105		110										
Thr	Ala	Thr	Cys	Trp	Cys	Val	Asn	Thr	Ala	Gly	Val	Arg	Arg	Thr	Asp
	115						120					125			
Lys	Asp	Thr	Glu	Ile	Thr	Cys	Ser	Glu	Arg	Val	Arg	Thr	Tyr	Trp	Ile
	130						135					140			
Ile	Ile	Glu	Leu	Lys	His	Lys	Glu	Arg	Glu	Ser	Pro	Tyr	Asp	Ser	Lys
145					150				155					160	
Ser	Leu	His	Thr	Ala	Leu	Gln	Lys	Glu	Ile	Thr	Thr	Arg	Tyr	Gln	Leu
			165					170						175	
Asp	Pro	Lys	Phe	Ile	Thr	Ser	Ile	Leu	Tyr	Glu	Asn	Asn	Val	Ile	Thr
	180							185					190		
Ile	Asp	Leu	Met	Gln	Asn	Ser	Ser	Gln	Lys	Thr	Gln	Asn	Asp	Val	Asp
	195						200					205			
Ile	Ala	Asp	Val	Ala	His	Tyr	Phe	Glu	Lys	Asp	Val	Lys	Gly	Glu	Ser
	210					215					220				
Leu	Phe	His	Ser	Ser	Lys	Lys	Met	Asp	Leu	Arg	Val	Asn	Gly	Glu	Gln
225					230					235				240	
Leu	Asp	Leu	Asp	Pro	Gly	Gln	Thr	Leu	Ile	Tyr	Tyr	Val	Asp	Arg	Asn
			245					250					255		
Ala	Pro	Glu	Phe	Ser	Met	Gln	Ala	Leu	Lys						
		260						265							

<210> 79

<211> 798

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA encoding Tag-25/18 chimera comprising
SP+PP+ECD

<400> 79

```

atggcaccac ctcaagcact ggcaactgggt ctgctgctgg ccgccgctac cgccactttc 60
gccgcagcac aggaggagtg tgtgtgcgaa aactacaagc tcgctgtcaa ctgtttcgtc 120
aacaataata gagaatgccg gtgcacttct gtgggagcac agaatacagt gatctgtagc 180
aaactggctg caaagtgtct ggcatgaag gccgaaatga acggatccaa gtcggggcgg 240
aggatcaaac ctgaaggagc tctgcagaac aacgatggct tctacgaccc cgactgtgac 300
gagtcgggcc tcttcaaggc caaacagtgt aatggcactg ctacatgctg gtgcgtgaac 360
accgctgggg tgcgccggac cgataaggat accgaaatta cctgttctga gaggggccgg 420
acatattgga tcatcattga actcaaact aaagagcgcg agtctccata cgattctaaa 480
tccctccata ctgcactgca aaaggaaatc actacacgct accagctgga tccaaaattc 540
attacatcca tcctctatga gaacaatgt attacaattg atctgatgca aaatagctct 600
cagaagactc agaatgatgt ggacatagct gatgtggctc attattttga aaaagatgtt 660
aaaggtgaat ccttgtttca ttcttctaag aaaatggacc tgagagtaaa tggagaacaa 720
ctggatctgg atcctggtca aactttaatt tattatgttg atagaaatgc acctgaattt 780
tcaatgcagg ctctaaaa

```

<210> 80

<211> 25

<212> PRT

<213> Artificial Sequence

<220>

<223> TAG-18 TMD

<400> 80

Ala Gly Val Cys Ala Val Ala Val Ile Val Val Val Met Ile Ala Val
 1 5 10 15
 Val Ala Gly Ile Val Val Leu Val Ile
 20 25

<210> 81
 <211> 53
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Seq Pattern 1 (propeptide alignment)

<220>
 <221> VARIANT
 <222> (1)...(53)
 <223> Xaa = any amino acid or as shown in Table 3

<400> 81
 Gln Xaa Xaa Cys Val Cys Xaa Asn Tyr Lys Leu Xaa Xaa Xaa Cys Xaa
 1 5 10 15
 Xaa Asn Xaa Xaa Xaa Xaa Cys Gln Cys Thr Ser Xaa Gly Xaa Gln Asn
 20 25 30
 Thr Val Ile Cys Ser Lys Leu Ala Xaa Met Lys Ala Glu Met Xaa Xaa
 35 40 45
 Ser Lys Xaa Gly Arg
 50

<210> 82
 <211> 53
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Seq pattern 2 - propeptide alignment with
 retention of epitopes

<220>
 <221> VARIANT
 <222> (1)...(53)
 <223> Xaa = any amino acid or as shown in Table 4

<400> 82
 Gln Xaa Xaa Cys Val Cys Glu Asn Tyr Lys Leu Ala Val Xaa Cys Xaa
 1 5 10 15
 Xaa Asn Xaa Xaa Xaa Xaa Cys Gln Cys Thr Ser Xaa Gly Xaa Gln Asn
 20 25 30
 Thr Val Ile Cys Ser Lys Leu Ala Val Met Lys Ala Glu Met Xaa Xaa
 35 40 45
 Ser Lys Xaa Gly Arg
 50

<210> 83
 <211> 23

<212> PRT
<213> Artificial Sequence

<220>
<223> Signal peptide sequence pattern

<220>
<221> VARIANT
<222> (1)...(23)
<223> Xaa = any amino acid or as shown in Table 5

<400> 83
Met Ala Xaa Pro Xaa Xaa Leu Ala Xaa Gly Leu Leu Leu Ala Xaa Xaa
1 5 10 15
Thr Ala Thr Xaa Ala Ala Ala
20

<210> 84
<211> 33
<212> PRT
<213> Artificial Sequence

<220>
<223> EGF-LIKE DOMAIN 1 SEQ PATTERN

<220>
<221> VARIANT
<222> (1)...(33)
<223> Xaa = any amino acid

<400> 84
Cys Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa
1 5 10 15
Xaa Xaa Xaa Cys Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
20 25 30
Cys

<210> 85
<211> 70
<212> PRT
<213> Artificial Sequence

<220>
<223> EGF-LIKE DOMAIN 2 SEQ PATTERN

<220>
<221> VARIANT
<222> (1)...(70)
<223> Xaa = any amino acid

<400> 85
Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

			20					25					30			
Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	
			35					40				45				
Xaa	Xaa	Cys	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	
		50					55					60				
Xaa	Xaa	Xaa	Xaa	Xaa	Cys											
65					70											

<210> 86
 <211> 33
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> EGF-LIKE DOMAIN 1 SEQ PATTERN with TAg-25 epitope

<220>
 <221> VARIANT
 <222> (1)...(33)
 <223> Xaa = any amino acid

Cys	Val	Cys	Glu	Asn	Tyr	Lys	Leu	Ala	Val	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	
1				5					10					15		
Xaa	Xaa	Xaa	Cys	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	
			20					25					30			
Cys																

<210> 87
 <211> 70
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> EGF-LIKE DOMAIN 2 SEQ PATTERN with TAg-25 epitopes
 and RR site

<220>
 <221> VARIANT
 <222> (1)...(70)
 <223> Xaa = any amino acid

<220>
 <221> VARIANT
 <222> (17)...(17)
 <223> Xaa = preferably Ala, Ile or Met

Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Arg	Arg	
1				5					10					15		
Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Gln	Asn	Asn	Asp	Gly	Leu	Tyr	Asp	Pro	
			20					25					30			
Asp	Cys	Asp	Glu	Ser	Gly	Leu	Phe	Lys	Xaa	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	
		35					40					45				

Ala Thr Cys Trp Cys Val Asn Thr Ala Xaa Xaa Xaa Xaa Xaa Xaa
 50 55 60
 Xaa Xaa Xaa Xaa Xaa Cys
 65 70

<210> 88
 <211> 109
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Combined EGF domains (all Cys) pattern

<220>
 <221> VARIANT
 <222> (1)...(109)
 <223> Xaa = any amino acid

<400> 88
 Cys Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa
 1 5 10 15
 Xaa Xaa Xaa Cys Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 20 25 30
 Cys Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 35 40 45
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 50 55 60
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 65 70 75 80
 Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Cys Xaa Cys Xaa Xaa Xaa Xaa
 85 90 95
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys
 100 105

<210> 89
 <211> 109
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Example of 12 Cys pattern with some epitopes

<220>
 <221> VARIANT
 <222> (1)...(109)
 <223> Xaa = any amino acid

<220>
 <221> VARIANT
 <222> (56)...(56)
 <223> Xaa = preferably Ala, Ile or Met

<400> 89
 Cys Val Cys Glu Asn Tyr Lys Leu Ala Val Xaa Cys Xaa Xaa Xaa Xaa
 1 5 10 15

Xaa	Xaa	Xaa	Cys	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
			20					25						30			
Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
		35					40						45				
Xaa	Xaa	Xaa	Xaa	Xaa	Arg	Arg	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Gln	Asn	
		50				55							60				
Asn	Asp	Gly	Leu	Tyr	Asp	Pro	Asp	Cys	Asp	Glu	Ser	Gly	Leu	Phe	Lys		
65					70					75					80		
Xaa	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	Ala	Thr	Cys	Trp	Cys	Val	Asn	Thr	Ala		
			85						90					95			
Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Cys				
			100					105									

<210> 90

<211> 55

<212> PRT

<213> Artificial Sequence

<220>

<223> Mature portion of SEQ ID NO:89

<220>

<221> VARIANT

<222> (1)...(55)

<223> Xaa = any amino acid

<220>

<221> VARIANT

<222> (2)...(2)

<223> Xaa = preferably Ala, Ile or Met

<400> 90

Arg	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Gln	Asn	Asn	Asp	Gly	Leu	Tyr	Asp		
1				5				10						15			
Pro	Asp	Cys	Asp	Glu	Ser	Gly	Leu	Phe	Lys	Xaa	Xaa	Xaa	Cys	Xaa	Xaa		
			20					25					30				
Xaa	Ala	Thr	Cys	Trp	Cys	Val	Asn	Thr	Ala	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa		
		35					40						45				
Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Cys											
		50				55											

<210> 91

<211> 71

<212> PRT

<213> Artificial Sequence

<220>

<223> Sequence pattern - Thyroglobulin type-1 repeat motif

<220>

<221> VARIANT

<222> (1)...(71)

<223> Xaa = any amino acid

<400> 91

Cys Xaa Val Glu Arg Xaa Xaa Xaa Xaa Xaa Xaa Ser Xaa Xaa Xaa Xaa
1 5 10 15
Xaa Xaa Xaa Xaa Glu Gly Ala Leu Xaa Xaa Xaa Xaa Gly Leu Tyr Xaa
20 25 30
Pro Xaa Cys Asp Glu Xaa Gly Xaa Xaa Lys Xaa Xaa Gln Cys Xaa Xaa
35 40 45
Xaa Xaa Xaa Xaa Cys Trp Cys Val Asp Xaa Xaa Gly Xaa Xaa Xaa Xaa
50 55 60
Xaa Xaa Asp Xaa Xaa Xaa Glu
65 70

<210> 92

<211> 184

<212> PRT

<213> Artificial Sequence

<220>

<223> TAG-25 fragment comprising ECD N-terminal variant

<400> 92

Ile Lys Pro Glu Gly Ala Leu Gln Asn Asn Asp Gly Leu Tyr Asp Pro
1 5 10 15
Asp Cys Asp Glu Ser Gly Leu Phe Lys Ala Lys Gln Cys Asn Gly Thr
20 25 30
Ala Thr Cys Trp Cys Val Asn Thr Ala Gly Val Arg Arg Thr Asp Lys
35 40 45
Asp Thr Glu Ile Thr Cys Ser Glu Arg Val Arg Thr Tyr Trp Ile Ile
50 55 60
Ile Glu Leu Lys His Lys Glu Arg Glu Ser Pro Tyr Asp Ser Lys Ser
65 70 75 80
Leu His Thr Ala Leu Gln Lys Glu Ile Thr Thr Arg Tyr Gln Leu Asp
85 90 95
Pro Lys Phe Ile Thr Ser Ile Leu Tyr Glu Asn Asn Val Ile Thr Ile
100 105 110
Asp Leu Met Gln Asn Ser Ser Gln Lys Thr Gln Asp Asp Val Asp Ile
115 120 125
Ala Asp Val Ala Tyr Tyr Phe Glu Lys Asp Val Lys Gly Glu Ser Leu
130 135 140
Phe His Ser Lys Lys Met Asp Leu Arg Val Asn Gly Glu Leu Leu Asp
145 150 155 160
Leu Asp Pro Gly Gln Thr Leu Ile Tyr Tyr Val Asp Glu Lys Ala Pro
165 170 175
Glu Phe Ser Met Gln Gly Leu Lys
180

<210> 93

<211> 795

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA encoding hEpCAM antigenic fragment comprising
SP+PP+ECD (termed "sEpCAM")

<400> 93

```
atggcgcccc cgcaggtcct cgcgttcggg cttctgcttg ccgcggcgac ggcgactttt 60
gccgcagctc aggaagaatg tgtctgtgaa aactacaagc tggccgtaaa ctgctttgtg 120
aataataatc gtcaatgcc a gtgtacttca gttgggtgcac aaaatactgt catttgctca 180
aagctggctg ccaaagtgtt ggtgatgaag gcagaaatga atggctcaaa acttgggaga 240
agagcaaaac ctgaaggggc cctccagaac aatgatgggc tttatgatcc tgactgcgat 300
gagagcgggc tctttaaggc caagcagtgc aacggcacct ccacgtgctg gtgtgtgaac 360
actgctgggg tcagaagaac agacaaggac actgaaataa cctgctctga gcgagtgaga 420
acctactgga tcatcattga actaaaacac aaagcaagag aaaaacctta tgatagtaaa 480
agtttgcgga ctgcacttca gaaggagatc acaacgcgtt atcaactgga tccaaaattt 540
atcacgagta ttttgtatga gaataatgtt atcactattg atctggttca aaattcttct 600
caaaaaactc agaatgatgt ggacatagct gatgtggctt attattttga aaaagatgtt 660
aaaggtgaat ccttgtttca ttctaagaaa atggacctga cagtaaatgg ggaacaactg 720
gatctggatc ctgggtcaaac tttaatttat tatgttgatg aaaaagcacc tgaattctca 780
atgcagggtc taaaa 795
```

<210> 94

<211> 372

<212> DNA

<213> Artificial Sequence

<220>

<223> nucleotide sequence of a mAb variable heavy chain domain

<400> 94

```
gaggtgaagc tgctggagtc cggaggtggc ctggtgcagc ctggaggatc cctgaaactc 60
tcctgtgcag cctcaggatt cgatttttagt agatactgga tgagttgggt ccggcagggt 120
ccagggaaag ggctagaatg gattggagat attaatctag aaagcaatac gataaactat 180
acgccatctc taaaggataa attcatcatc tccagagaca acgcaaaaaa tacgctgtac 240
ctgcaaatga acaaagttag atctgaggac acagcccttt attactgtgc aagagggggc 300
tatactatgg actactgggg tcaaggaacc tcagtcaccg tctcctcagc caaacgcaca 360
cccccatctg tc 372
```

<210> 95

<211> 126

<212> PRT

<213> Artificial Sequence

<220>

<223> amino acid sequence of a mAb variable heavy chain domain

<400> 95

```
Glu Val Lys Leu Leu Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
 1           5           10           15
Ser Leu Lys Leu Ser Cys Ala Ala Ser Gly Phe Asp Phe Ser Arg Tyr
 20           25           30
Trp Met Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Ile
 35           40           45
Gly Asp Ile Asn Leu Glu Ser Asn Thr Ile Asn Tyr Thr Pro Ser Leu
 50           55           60
Lys Asp Lys Phe Ile Ile Ser Arg Asp Asn Ala Lys Asn Thr Leu Tyr
 65           70           75           80
Leu Gln Met Asn Lys Val Arg Ser Glu Asp Thr Ala Leu Tyr Tyr Cys
 85           90           95
Ala Arg Gly Ala Tyr Thr Met Asp Tyr Trp Gly Gln Gly Thr Ser Val
```


			100				105			110		
Thr	Val	Ser	Ser	Ala	Lys	Thr	Thr	Pro	Pro	Ser	Val	Ala
		115					120					125